Bandolier 111

Independent evidence-based health care

Making too much from too little

One of the most difficult areas in our world of evidence is when there is not much information, and we have to make the best of what we have. Our sin, all too often, is that we make too much of too little. It **is** hard to stop ourselves from doing it, and the trouble is that, if we are enthusiastic believers, we can convince others of our case when the evidence just isn't there.

The Scottish legal system has a verdict that is neither guilty, nor innocent, but rather "not proven". It is rarely used and indicates when there is insufficient evidence to convict, but rather too much to be certain of innocence. It is a verdict that we might use profitably when thinking about evidence in our own field.

Making too much of too little is a sin too frequently encountered in systematic reviews, and one where "not proven" is much more satisfactory. ***Bandolier*** will return to this theme, but this month offers two examples. One is counselling in primary care, where the evidence from trials is anything but overwhelming. Some may consider that evidence from practice is good, but where is it? Another, the use of finasteride for haematuria in BPH where other causes have been excluded, shows that limited evidence from trials accords with limited evidence from practice.

On diagnostic tests

Readers have asked why ***Bandolier*** seems to have withdrawn from examining evidence about diagnostic tests. The answer was simply that we were having trouble finding better ways to describe the results, and even likelihood ratios were insufficiently intuitive for some. Using natural frequencies, and probability of disease with a positive or negative test, is one possible way forward.

Two examples are offered. One comes from a systematic review of studies of the Ottawa ankle and foot tests, clinical paradigms designed to exclude fractures and save on radiographs. Over a decade we now have lots of evidence, and the results are interesting. Another involves the results of the quadruple test for Down's syndrome screening.

Bandolier would like to move this forward, so any comments would be helpful, as would news of any studies on diagnostic tests readers (including industrial readers) would like to bring to our attention.

Counselling in primary care

The place, if any, of counselling in primary care is one of those topics that continues to attract attention. There is a Cochrane review, and some countries, including the UK, have produced guidelines based on evidence. What is the evidence, and how good is it? An updated review [1] by the authors of the Cochrane review is helpful.

Review

The search strategy for the original Cochrane review was heroic, and included 10 electronic databases, handsearching specialist journals, as well as consulting experts. The updated review used six databases, including specialist trials registers. Randomised and controlled trials were eligible that tested the hypothesis that counsellors treating patients in primary care are more effective than usual care provided by the GP or alternative mental health treatments.

Participants were patients consulting a GP with psychological or psychosocial problems. Intervention by a counsellor was undefined other than counsellors had to have been trained to the British standard for accreditation. Outcomes were clinical effectiveness of psychological outcomes like depressive symptoms and measures of social function.

Results

Seven trials were included from 12 publications, with 444 patients given counselling and 297 given usual care. On a quality scoring system all scored reasonably well, but how many of these trials were randomised or used blinded outcome assessments is not mentioned in this review.

Six trials had short-term outcomes (six weeks to six months). Three, including two with shorter observations, showed

In this issue

significant benefit of counselling over usual care, and three, all with longer observations, did not. The overall outcome was statistically significant, but of a modest size. Results in terms of outcomes for patients are best described by Figure 1. With counselling there was a small increase in the percentage of patients having a reliable and clinically meaningful change.

Four trials had longer-term outcomes (9-12 months). There was no significant difference between counselling and usual care.

Comment

This is an interesting and thoughtful review, well worth reading if primary care counselling is of interest. There are sensitivity analyses, and useful discussion. The authors do not seek to disguise the basic problem, that of insufficient information. Some of the issues are worth looking at, because they arise time and again in meta-analysis:

- Who were the patients? Some were depressed, some anxious, some had relationship or family problems, some were bereaved, some had sexual difficulties, or substance misuse problems. Some trials used a mix of patients, others used just patients with depression, or anxiety, or emotional problems.

- What was the counselling intervention? It was not always described, and rarely standardised. The number of sessions varied, and not all patients attended.

- What was usual care? Usually it was not described in the papers to any satisfactory degree, other than specifying that counsellors were trained to a British standard.

- What was the outcome? Good question, this, as ***Bandolier*** has no idea. The best came from descriptions of change (Figure 1). But describing this is very, very, difficult.

- How good were the trials? Pretty good using a specialist scoring system, but we do not know whether they were all randomised, or whether a blinded assessor made assessment of outcome. Without this knowledge we have no idea about potential sources of bias. A question that needs answering was the gross imbalance in numbers in some trials.

- How much information do we have? At best on 444 patients given counselling in comparisons with usual care. The important result, that of a difference in the proportion of patients with reliable and statistically significant change, depended on just 108 patients.

And yet this is the only information we have that can inform the question of counselling in primary care. Who among us would conclude either that it works, or does not work? The best response is that we cannot possibly know, but that large advantages of counselling are unlikely. Saying any more is to make far too much from far too little.

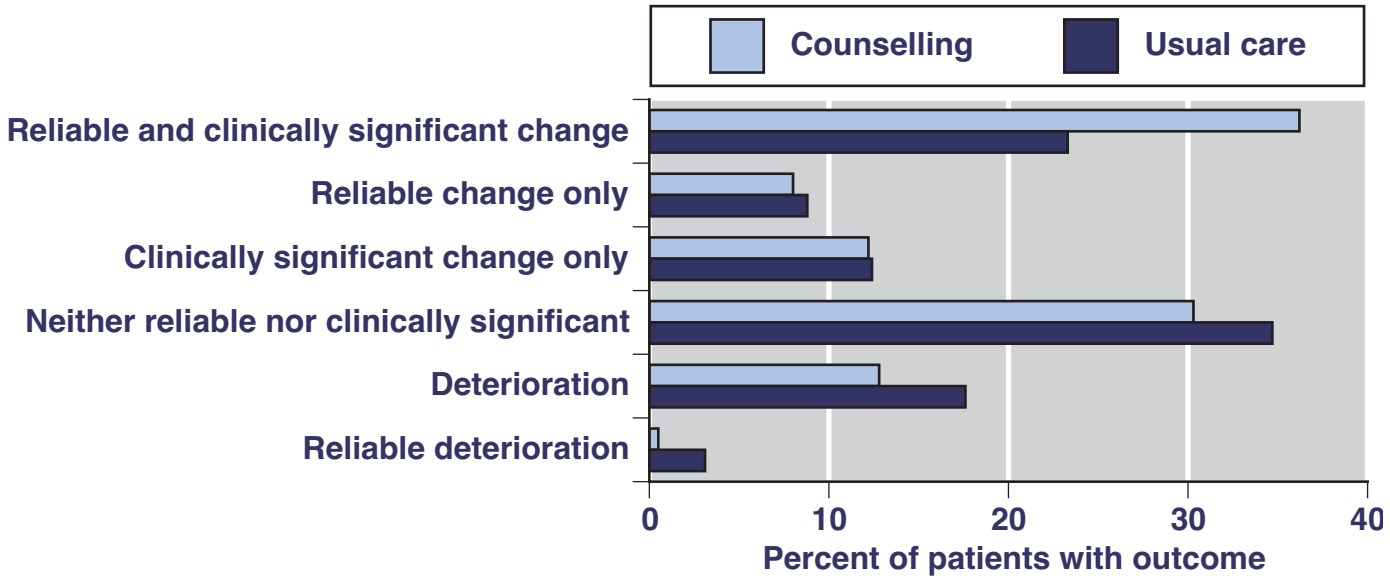
It is relevant to compare the weight and quality of evidence we have here with the weight and quality of evidence we expect from a newly introduced pharmacological therapy. There is little comparison. It is not even possible to say that counselling is better than usual care, and the trials say nothing about possible harms. For instance, might there be rare but serious harm from counselling that outweighs any possible small benefit?

Again, it is not possible to say anything about any cost consequences, because without knowing anything about effectiveness, we can say nothing about costs. On the basis of the evidence we have from this review, would it be a sensible decision to begin a widespread use of counselling in primary care?

References:

1 P Bower et al. The clinical effectiveness of counselling in primary care: a systematic review and meta-analysis. Psychological Medicine 2003 33: 203-215.

Figure 1: Outcomes in counselling and usual care in a mixed population of patients with different problems in primary care



BPH, HAEMATURIA, AND FINASTERIDE

A question asked of ***Bandolier*** has been whether treating haematuria associated with benign prostatic hyperplasia (BPH) with finasteride works or is worthwhile. This has not been easy to answer because there have been few randomised trials, and few observational studies that address the topic. But recent months have seen some new studies that now allow a sensible look at the evidence.

Search

We searched PubMed, the Cochrane Library and reference lists up to February 2003 for any studies relating to haematuria (or hematuria), BPH, and finasteride. For the purposes of this brief review studies of any architecture were permitted, though randomised trials were preferred. All studies whose abstracts appeared to be relevant were obtained and read. Outcomes sought were recurrence of haematuria, any prostate surgery, and adverse events.

Results

There were three randomised trials and seven prospective or retrospective observational studies. Three observational studies related to a similar patient set, and the one with the most useful information was used to avoid duplication. Details of the eight studies included are in Table 1 (page 4). The dose of finasteride, where stated, was 5 mg a day.

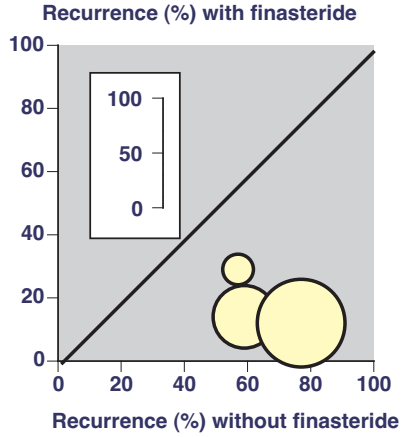
Men in the studies were elderly, with mean ages in the 70s, and all had BPH with clear exclusion of any other cause of their haematuria, including PSA or biopsy for cancer.

Randomised studies

The three randomised studies were small and open, and none gave details of randomisation or concealment. Quality scores were 2 out of a possible 5 on a commonly used scale, indicating the possibility of bias. The largest study was conducted over 48 months, and the others over 12 months. The control was no treatment or watchful waiting.

Recurrence of haematuria to any degree occurred in 14/92 (15%) men on finasteride compared with 48/73 (66%) with no treatment (Figure 1). The risk of recurrence of haematuria was reduced by 75% with finasteride (Table 2), and the number needed to treat (NNT) was 2 (1.6 to 2.7) to prevent recurrence. Two studies noted the severity of recurrent haematuria, and it was mostly a single episode of bleeding and of minor severity.

Figure 1: Recurrence of haematuria with finasteride and placebo



The necessity for prostate surgery (almost always transurethral prostatectomy) was reduced with finasteride, occurring in 5/92 (5%) of men on finasteride and 16/73 (22%) with no treatment. The risk of prostate surgery was reduced by 78% with finasteride (Table 2), and the number needed to treat (NNT) was 6 (3.7 to 17) to prevent surgery. In one study three other men having no treatment required hospital admission for cystoscopy.

Observational studies

The five observational studies examined cohorts of men with haematuria and BPH for various times, and reported the recurrence of haematuria (Table 1). The inclusion criteria were similar to men in the randomised trials.

Finasteride was given to 120 men, and 19 (16%) of those had recurrence of haematuria. In some, recurrence was associated with intermittent use of finasteride. No TURP surgery was reported in 53 men in two reports.

Comment

The amount and quality of information is neither large nor convincing, yet the results are consistent. Both randomised and observational studies yielded recurrence rates of 15% or 16% with finasteride, much less than was seen with no treatment in the randomised trials. NNTs were low. For every two men with BPH and haematuria with no other cause treated with finasteride one recurrence would be prevented, and for every six men one TURP avoided.

Treating a man with finasteride for a year is not cheap, at about £300. But recurrence has costs of possible cystoscopy, and prostatectomy has both costs and clinical problems in men in their late 70s. Formal cost-benefit analysis may be premature, but this certainly looks like an effective treatment in this particular group of men.

Table 2: Outcomes in randomised studies, preventing recurrence of haematuria and TURP

Outcome	Percent with finasteride	Percent without finasteride	Relative risk (95%CI)	NNT (95%CI)
Prevent recurrence of haematuria	15	66	0.24 (0.15 to 0.40)	2.0 (1.6 to 2.7)
Prevent TURP	5	22	0.22 (0.08 to 0.58)	6.1 (3.7 to 17)

Table 1: Finasteride for haematuria in BPH - randomised and observational studies

Reference	Design	Included men	Outcome	Results
Foley et al. J Urol 2000 163: 496-498	Randomised, open study of 57 men with watchful waiting or finasteride 5 mg for 12 months	Clinically documented BPH presenting with gross haematuria with no other cause. Average age 78 years, some with previous surgery	Bleeding episodes, with severity	Over 12 months: watchful waiting: 17/29, 7 minor, 6 moderate, 4 severe finasteride: 4/28, 3 minor, 1 moderate 7 hospital admissions in controls, 4 TURP, 3 cystoscopy
Delakas et al. Urol Int 2001 67: 69-72	Randomised, open study with 50 men treated with 5 mg finasteride daily for 4 years, and 30 controls (watchful waiting).	Clinically documented BPH (after TURP in 17). At least 4 episodes of haematuria, mean age 74 years	Recurrence of haematuria and surgery	At 48 months: control: recurrence 23/30 TURP 9/30 finasteride: recurrence 6/50 TURP 4/50
Perimenis et al. Urology 2002 59: 373-377	Randomised, open study of 42 men with watchful waiting, finasteride 5 mg, or cyproterone acetate 100 mg (daily doses) for 12 months	Clinically documented BPH presenting with gross haematuria with no other cause. Average age 77 years,	Bleeding episodes, with severity	Over 12 months: watchful waiting: 8/14, 2 minor, 2 moderate, 4 severe (2 TURP) finasteride: 4/14, 2 minor, 2 moderate (0 TURP) CYA: 3/14, 2 minor 1 moderate (0 TURP) No adverse event discontinuations, and some impotence with CYA
Carlin et al. Prostate 1997 31: 180-182	Prospective observational study of 12 men treated with 5 mg finasteride daily for 6 months	Gross haematuria secondary to BPH, excluding other causes	Bleeding	Bleeding subsided in 2 weeks in all patients, with no recurrences over mean of 11 months
Kashif et al. Prostate Cancer & Prostatic Diseases 1998 1: 154-156.	Retrospective review of 42 men at haematuria clinic	Bleeding associated with BPH and no other cause, 52-89 years, some with previous surgery	Recurrence of haematuria and surgery	No treatment, 6-18 months: 6/18 2 or more episodes 10/18 no episodes 3 TURP Finasteride, 4-21 months 0/24 2 or more episodes 22/24 no episodes 0/24 TURP two patients with minor adverse events with finasteride
Sieber et al. J Urol 1998 159: 1232-1233	Retrospective review of 42 men treated with finasteride at haematuria clinic	Bleeding associated with BPH and no other cause, average age 74 years, some with previous surgery	Recurrence of haematuria	Over 6-40 months in 28 evaluable patients: 2/28 minor bleeds 1/28 severe bleeds 25/28 no bleeds
Miller & Puchner. Urology 1998 51: 237-240	Retrospective review of 28 men treated with finasteride for haematuria	Bleeding associated with BPH and no other cause, age 62-87 years, some with previous surgery, some on aspirin or warfarin	Recurrence of haematuria	Over4-47 months: 5/28 minor bleeds 2/28 moderate bleeds 1/28 severe bleeds (intermittent use) 18/28 no bleeds 2 discontinued treatment
Redorta et al. Arch Esp Urol 2002 55: 895-899	Prospective observational study of 29 men treated with 5 mg finasteride daily for an average of 15 months	Haematuria from BPH, mean age 71 years, many with previous prostate surgery	Recurrence of haematuria, and surgery	25/29 no recurrence 0/29 TURP

Ottawa ankle rules revisited

As long ago as ***Bandolier*** 21 the Ottawa ankle rules were featured in these pages as an example of how clinical diagnostic tests can be assessed and evaluated. Since then the rules have been evaluated in numerous studies, so that we now have a meta-analysis [1]. ***Bandolier*** also thought that this might be a useful opportunity to contrast traditional methods of describing diagnostic tests (sensitivity, specificity, likelihood ratios) with the natural frequency methods suggested by Gerd Gigerenzer (***Bandolier*** 109).

Review

The review used a systematic search for studies of the Ottawa ankle and foot rules, using several databases and without language restriction. For each study information was sought on methodological issues, such as whether enrolment was consecutive, whether radiologists were blinded, and whether radiography was used in all patients. Pooled assessment was made for sensitivity, but not specificity, and negative likelihood ratios calculated, with sensitivity analyses.

Results

In total 32 studies were found, some looking at the ankle rules, some at the foot rules, some at both, and while most were in adults, some were in children. In 27 studies with data for a pooled analysis, out of 15,581 patients, 27 (0.3%) had a false negative result where the Ottawa test was negative, but where they actually had a fracture.

Overall the pooled sensitivity (percentage with a fracture testing positive and correctly classified as such) was 98%, and most studies achieved very high levels of sensitivity. Specificity (the percentage without a fracture who tested negative) was highly variable, some studies being as low as about 10%, with most at about 40%, and some as high as about 70%.

The likelihood ratio for a negative test was about 0.1, meaning that with a fracture prevalence of about 10%, the chance of there actually being a fracture was about 1 in 100.

Comment

The Ottawa ankle and foot rules were designed to minimise the number of radiographs needed. Specificity was the key to this. Table 1 shows the calculations for a cohort of 1,000 persons, in whom there were 100 actual fractures, applied to the best, average, and worse specificity values found in the review. As specificity declined, many more positive tests, most of them false, would be found, requiring more radiographs. As specificity declines, the reason for the clinical diagnostic test diminishes. The ideal specificity of about 0.9 would yield only about 200 tests out of 1000 people. The actual results required between 350 and 900.

With the best the probability of a fracture with a negative test was 1 in 330, and with a positive test was 1 in 4. High probabilities are best with a negative test, and low probability best for a positive test. The ideal result would mean that 1 in 2 people with a positive test would actually have a fracture, and only 1 in 400 with a negative test.

Figure 1 shows how this looks for the example of the best specificity found in the review, using natural frequencies. It allows the calculation of these probabilities rather easily. Now ***Bandolier*** has always had problems with sensitivity, specificity, and likelihood ratio definitions. Each time we come to them we open David Sackett's books and start from scratch, and have to use our bespoke Excel sheet to do the calculations. And even then it needs several cups of strong coffee and some aching neurones before we get it. Natural frequencies seem easier to ***Bandolier***, and the output, of chance of disease with positive or negative tests, seems intuitive and useful. Worth persevering with.

References:
1 LM Bachmann et al. Accuracy of Ottawa ankle rules to exclude fractures of the ankle and mid-foot: a systematic review. BMJ 2003 326: 417-423.

Figure 1: Ottawa ankle and foot rules using natural frequencies and "best" results for specificity from systematic review

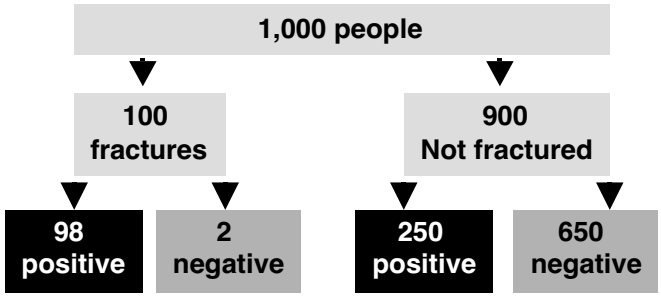


Table 1: Results of studies of Ottawa ankle and foot rules, using an ideal scenario, and best, average, and worst specificities from systematic review

			Number of tests		Chance of a fracture in test	
Scenario	Sensitivity	Specificity	Positive	Negative	Positive	Negative
Ideal specificity	0.98	0.9	198	802	1 in 2	1 in 400
Best specificity	0.98	0.7	348	652	1 in 4	1 in 326
Average specificity	0.98	0.4	648	352	1 in 7	1 in 176
Worst specificity	0.98	0.1	898	102	1 in 9	1 in 51

Outcomes predicted from a cohort of 1000 people presenting with possible fractured ankle, in which 100 (10%) actually have a broken ankle

Risk and Down's screening

Screening for Down's syndrome with four measurements in maternal serum is increasingly common. The four markers are alphafetoprotein, unconjugated oestriol, human chorionic gonadotrophin and inhibin-A. Results are used by a computer programme, together with gestational age, to compute a risk of an affected child. A positive test is one in which the risk of an affected child is higher than 1 in 300. We now know how this worked in practice in nearly 50,000 women over five years [1].

Study

The population were 46,193 pregnancies in 14 London hospitals over five years, in which the quadruple test was applied to serum samples between 14-22 weeks of pregnancy. A test was deemed positive if the computed risk of an affected foetus was 1 in 300 or greater (where greater risk means lower numerical values, 1 in 200, 1 in 100 etc). Gestational age was determined by ultrasound in 4 of 5 women.

Results

There were 88 affected pregnancies, giving an overall risk in this population of 1 in 525, not taking age into account. With the quadruple test there were 3,271 positive tests that detected 71 affected foetuses. Just under 98% of all positive tests were false positives. Sensitivity, specificity and likelihood ratios are shown in Table 1. The chance of an affected foetus following a positive test was 1 in 46, and with a negative test was 1 in 2525.

By comparison, maternal age alone, using 35 years or older as a cut off, would have detected 45 affected foetuses had 6,659 women proceeded to amniocentesis. The chance of an affected foetus using age over 35 alone was 1 in 148, but was 1 in 850 with a negative test.

Comment

This paper also has information on the use of triple and double tests (where inhibin-A, and inhibin-A and unconjugated oestriol were not included, respectively), but that information was not amenable to similar calculations. The paper also tells us that the uptake of amniocentesis rose with increasing risk, from 43% of women with risks between 1 in 250 to 1 in 300, to 74% in those with risks higher than 1 in 50.

Of women who tested positive and had an affected pregnancy, 62 of 71 had amniocentesis, and 59 of the 62 had a termination. Twenty children were born with Down's syndrome.

Natural frequencies for the quadruple test and maternal age are shown in Figure 1.

To calculate the chance of an affected foetus with a positive test, the 71 actual cases detected are divided into the sum of all positive tests, in the case of the quadruple test:

$$(3{,}200 + 71)/71 = 46 \text{ (1 chance in 46)}$$

To calculate the chance of an affected foetus with a negative test, the 17 cases not detected are divided into the sum of all negative tests, in the case of the quadruple test:

$$(42{,}905 + 17)/17 = 2525 \text{ (1 chance in 2525)}$$

References:
1 NJ Wald et al. Antenatal screening for Down's syndrome with the quadruple test. Lancet 2003 361: 835-836.

Figure 1: Screening for Down's syndrome using natural frequencies for the quadruple test and maternal age

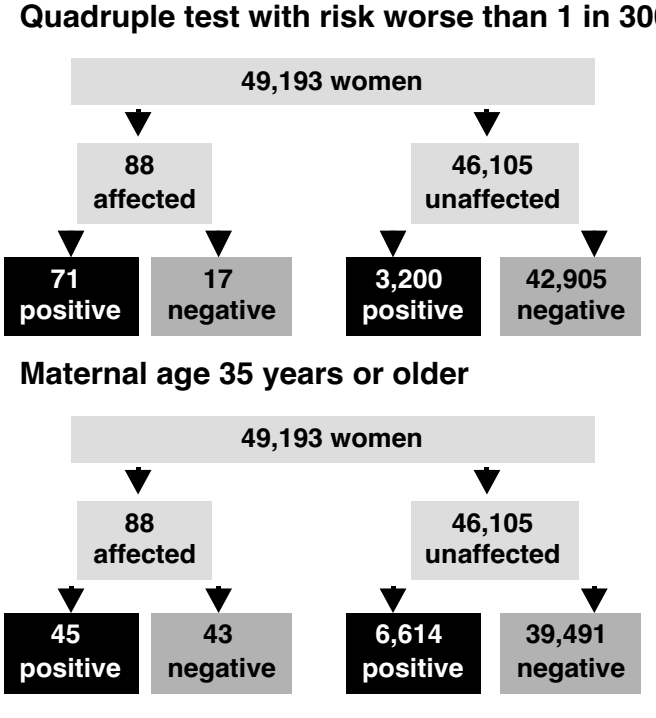


Table 1: Results of Down's syndrome screening using the quadruple test and maternal age

			Likelihood ratio		Chance of an affected foetus with test	
Scenario	Sensitivity	Specificity	Positive	Negative	Positive	Negative
Quadruple test	0.81	0.93	11.6	0.21	1 in 46	1 in 2525
Maternal age 35 or older	0.51	0.79	2.5	0.62	1 in 148	1 in 850

Outcomes predicted from a cohort of 46,193 women tested over 5 years at London hospitals. The risk of an affected foetus overall was 1 in 525.

Handwashing in schoolchildren

Bandolier 91 reported that encouraging handwashing in military recruits reduced outpatient visits for respiratory conditions. Another target for improving washing of hands is schoolchildren, and a study in Detroit [1] showed handwashing to be effective in reducing absence through illness in elementary school children.

Study

A single school with 305 children aged 5-12 years in 14 classes was chosen, and classes were divided into experimental and control without formal randomisation. Six classrooms with 143 children formed the experimental group and eight with 162 children the control group.

The intervention was that children in the handwashing classes were required by their teachers to wash their hands after arrival at school, before eating lunch, after the lunch break and before going home. This was done as a class activity. Children in the control group had no required hand washing.

The outcome was absence from school, recorded daily and monitored by telephone contact with parents to investigate the nature of any illness. Illness was regarded as respiratory if it included cough, sneeze, sinus trouble etc, and gastrointestinal if it included abdominal pain or diarrhoea or vomiting. Both were used, and total days of absence for illness. Children with both respiratory and gastrointestinal reasons for absence were included in both groups, so the total may be less than the sum of these two reasons.

Results

Absences were recorded as a percentage of the total number of possible days of attendance. Absences were lower in the handwashing than in the control group (Figure 1). The relative risk for each reason is given in Table 1. For total illness absence and gastrointestinal reasons, the reduction was statistically significant.

Figure 1: Days absent, total and by cause, for children washing hands regularly, and control

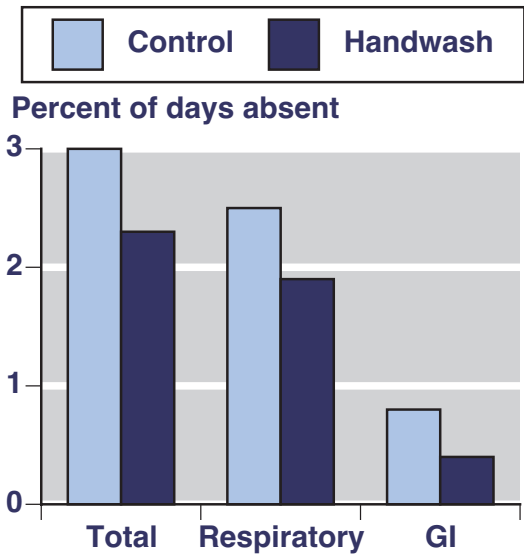


Table 1: Relative risks for absence, total and by cause, for handwashing compared with control

	Relative risk (95% CI)
All illness	0.75 (0.60 to 0.95)
Respiratory illness	0.79 (0.61 to 1.02)
Gastrointestinal illness	0.43 (0.25 to 0.73)

Comment

This study was not randomised, nor was it blind. Both these defects could have contributed to a result that might be biased. But there is so little evidence for the effectiveness of handwashing in the community that this study is worth a look. It also reports that it performed a literature review, though no details are given. There are perhaps three other studies in children, usually in special circumstances. Mostly they showed reduction in illness, and we know that in other settings washing hands frequently makes a difference.

What is surprising is that there is so little information about so basic an activity. Where are the trials, and where are the reviews of trials that can help us determine how and where washing hands can help, both in the community, as here with schoolchildren, and in healthcare to reduce infections?

Bandolier has started a special handwashing section on its Internet pages, to begin to draw together the evidence we have about handwashing. Searching electronic databases reveals few trials. Perhaps it is because there are few academic brownie points in so simple an issue.

Readers could help by telling *Bandolier* about studies they know of, and email or fax references, or better still the papers if they are in uncommon journals. We would also like to hear from those who have successfully implemented handwashing schemes.

References:

1 D Master et al. Scheduled hand washing in an elementary school population. Family Medicine 1997 29: 336-339.

Subscribing to Bandolier

Readers who would like their own copy have several ways of subscribing. You can:

Download a subscription form from the Internet (www.jr2.ox.ac.uk/bandolier/backnos.html) and fax it to Maura Moore on +44 1865 226978

Contact Maura for a subscription form by telephone (+44 1865 226132), fax (+44 1865 226978), or by email to maura.moore@pru.ox.ac.uk

Subscription rates are £36 in the UK and £72 overseas.

Book reviews

Introducing palliative care. Robert Twycross. Radcliffe Medical Press 2003, Fourth Edition. ISBN 1-85775-915-X, 190 pp, £not known.

Palliative care is an interesting topic for evidence-based medicine. First of all, there is little evidence in terms of randomised controlled trials, and therefore few systematic reviews, though the Pain, Palliative, and Supportive Care Cochrane review group and their reviewers are beginning to help us understand what evidence there is. Then there is the question of the difficulty of conducting clinical research in palliative care.

It is complicated by palliative care being a complex package of care tailored to each individual patient, and the fact that, by its very nature, patients die, making conventional trials difficult. After a while one wonders about whether this makes clinical trials themselves the best way to understand evidence, and though that is forced by the difficulties, satisfactory alternatives are themselves hard to find.

In these circumstances, the insights from a lifetime of experience help, and Robert Twycross provides this in the most recent edition of his book. There is lots of good stuff here about things that might not always be in the forefront of one's mind, like the special requirements of different religious beliefs. But perhaps what sets it out as special is the section on symptom management, with evaluation, explanation, management, monitoring and attention to detail (EEMMA as an acronym), and particularly the guidelines. These include guidelines for starting a patient on oral morphine, or transdermal fentanyl, or management of opioid constipation.

All in all a useful book for professional carers, and an opportunity for the younger and less experienced to gain some valuable insights into palliative care.

Systematic reviews to support evidence-based medicine. Khalid Khan, Regina Kunz, Jos Kleijnen, Gerd Antes. Royal Society of Medicine Press 2003. ISBN1-85315-525-X. 135 pp, £not known.

If you want an unashamedly Cochrane view of systematic reviews, then this is the book for you. It sets out the steps of framing questions, identifying literature, assessing quality, summarising evidence and interpreting findings in exactly the way it is done by the Cochrane Collaboration. It is one way of looking at the world of reviews, but not everyone will agree with it.

Bandolier readers will find it unfamiliar territory, because they will find few L'Abbé plots or NNTs, and instead be immersed in Forrest plots, odds ratios and relative risk. There is much talk of heterogeneity, and the material on clinical heterogeneity is good. But there is also stuff on funnel plots for publication bias that may just be plain wrong.

Validity of trials is a difficult, but enormously important topic. Validity can mean many things, but the single most important meaning is to that of whether a trial has the ability to answer a question. The dictionary definition is that which is sound or defensible. Validity is situation dependent, and criteria for validity might include the severity of a condition, the dose or intensity of intervention, the duration of an intervention or the duration of observation. Validity is crucial to what trials go into a review. Yet validity is found only as a glossary entry in the book.

This book will undoubtedly help some people to do their own reviews, and some of those reviews will be right. The worry is whether the book will stop them doing reviews which are wrong. That may be too harsh a judgement, but it is a worry.

Medical statistics made clear. Ashis Bannerjee. Royal Society of Medicine Press 2003. ISBN 1-85315-544-6. 137 pp, £not known.

So what is a z score? Or a risk ratio? Or a sensitivity analysis? Most of us have problems with some or all of these. We sort of know what they mean, but would be hard pressed to write an essay on them, and we hope we understand them enough to make sense of other people's use of them.

This neat little book is not a statistical text book, but rather a series of definitions and explanations, mostly in layman's language, and without requiring any great mathematical knowledge or expertise. The section on epidemiological concepts has interesting sections listing things like evidence supporting causality, and evaluation of randomised trials, and that old bugbear, confounding. There is even a nice definition of the Hawthorne effect. Later chapters even have simple things well explained, like a Latin square.

This is a useful little book for the shelf, and you will soon find yourself thumbing through it.

The menopause. What you need to know. Margaret Rees, David Purdie, Sally Hope. BMS Publications 2003. ISBN 0-9536228-2-7. 105 pp £6.99.

The updated 2002 version of this useful monograph, which has been around since 1994. Bandolier reviewed its predecessor in 1999 and found it useful then. It is even better now, with some nice treatment of risks. As before, it is a useful book for healthcare professionals, but will, notwithstanding, be of interest to informed consumers, so acessible is the writing.

EDITORS

Andrew Moore Henry McQuay
Pain Relief Unit
The Churchill, Oxford OX3 7LJ

Editorial office: 01865 226132
Editorial fax: 01865 226978
Email: andrew.moore@pru.ox.ac.uk
Internet: www.ebandolier.com
ISSN 1353-9906